

Open Research Online

The Open University's repository of research publications and other research outputs

State of the App: A Taxonomy and Framework for Evaluating Language Learning Mobile Applications

Journal Item

How to cite:

Rosell-Aguilar, Fernando (2017). State of the App: A Taxonomy and Framework for Evaluating Language Learning Mobile Applications. CALICO Journal, 34(2) pp. 243–258.

For guidance on citations see [FAQs](#).

© 2017 Equinox Publishing



<https://creativecommons.org/licenses/by-nc-nd/4.0/>

Version: Accepted Manuscript

Link(s) to article on publisher's website:

<http://dx.doi.org/doi:10.1558/cj.27623>

Copyright and Moral Rights for the articles on this site are retained by the individual authors and/or other copyright owners. For more information on Open Research Online's data [policy](#) on reuse of materials please consult the policies page.

oro.open.ac.uk

State of the App: a taxonomy and framework for evaluating language learning mobile applications

Fernando Rosell-Aguilar

School of Languages and Applied Linguistics

The Open University

Fernando Rosell-Aguilar is a Senior Lecturer in Spanish and Open Media Fellow at the Open University, United Kingdom. His research focuses on online language learning, mainly the use of apps, Twitter, and podcasting as teaching and learning tools, as well as the use of CMC learning environments (such as audio and video conferencing) and digital literacies.

Abstract

The widespread growth in availability and use of smartphones and tablets has facilitated an unprecedented avalanche of new software applications with language learning and teaching capabilities. However, little has been published in terms of effective design and evaluation of language learning apps. This paper reviews current research about the potential of apps for language learning and presents a taxonomy of available apps and their use for language learning. The paper also presents a framework consisting of four categories for evaluating language learning apps (technology, pedagogy, user experience, and language learning) and a set of criteria within the categories. Finally, the paper proposes areas for further research.

Keywords: Apps; Language Learning; Taxonomy; Evaluation; Framework

1 Introduction

The market penetration of smartphones and tablets has been very fast and widespread. The impact of these devices is due in part to features that at the time of launch were either new or vast improvements on previous mobile phones, including larger screen size, responsive touch screen, enhanced text-entry, high-quality audio and video playback, recording and editing, voice recognition, enlarged storage, and faster connectivity (Godwin-Jones, 2011). Other features include portability and intuitive interfaces.

Before 2007 most mobile phones only carried the software provided with the device, but this changed with smartphones as they included the ability to add additional software applications. These software applications for mobile devices are commonly known as apps (or mobile apps). Apps can be downloaded from app stores for different operating systems, which offer a category of apps named Education, with apps aimed at wide-ranging learning subjects, including languages. Many Apps can be downloaded for free, whereas others need to be paid for, usually at quite a low cost. Some apps offer a free “lite” version of the app so users can try them and decide whether to buy the full version and other apps offer in-app purchases to access further content or remove advertisements.

2 Apps for language learning: a literature review

2.1 Potential, criticisms, previous studies and taxonomies

The availability of apps has provided affordances for educational activity in terms of what can be done, where and when, with a single device. Among the potential advantages first identified for language teaching and learning were the opportunities to teach, practice or enhance a number of language learning skills as well as learners’ knowledge of the areas where the target language is spoken (Rosell-Aguilar, 2009). Other authors have further highlighted the potential of smartphone and tablet devices, as well as apps, for language learning (Burston, 2014; Godwin-Jones, 2011; Kim & Kwon, 2012; Kim, 2013; Lys, 2013; Sweeney & Moore, 2012). This potential is based on the theoretical principles and evidence from the field of Mobile-Assisted Language Learning (MALL). Among these are the provision of resources that can be used autonomously, taking screen size into consideration in the design of resources, and chunking knowledge as independent learning objects to facilitate processing of information (Ally, 2005). Other principles that apply to the use of mobile apps for language teaching are from the field of gamification, the use of game design elements in educational contexts (Domínguez et al. 2013).

App design for language learning has come under criticism: Burston (2014) argued that language learning activities on mobile apps are basic and have mostly replicated what was done before with other technologies. Although most practitioners in Computer-Assisted Language Learning (CALL) would agree that design for online language learning and teaching should be pedagogy-driven (Colpaert, 2006), many language learning apps often provide exercises that test the user without providing teaching first, or provide only a few very brief examples of use. In addition, feedback on performance tends to be limited to a tick or a cross to indicate whether an answer is correct or incorrect. They also tend to lack full instructions and their help sections, if at all available, address technical rather than pedagogical issues. Further criticisms related to the design of language learning apps include too much focus on translation, poor navigation and user-interface design, and little use of the unique properties of smartphones - connectivity with other users in particular (Godwin-Jones, 2011; Burston, 2014).

Other researchers agree: in their review of language learning mobile apps, Kim and Kwon (2012) highlighted that most apps focus on cognitive processes (recognition, recall and comprehension) and receptive language skills. They note the lack of socio-cognitive activities or opportunities for collaborative learning, more consistent with more modern approaches to CALL and MALL. What a CALL practitioner considers good practice, however, may not be what users want. As classroom practice has moved towards more modern approaches, learners may feel the need for more grammatical reinforcement in the form of drilling, given that many learners equate learning a language with learning grammar. Since individualized feedback on performance is something many learners rarely get outside formal tuition, getting answers correct in quizzes, or using apps to memorize verb forms and vocabulary, are rewarding activities and users are afforded the satisfaction of knowing they got something right. Whilst some apps continue to offer drilling with little teaching and lack of meaningful

feedback or support, some examples of good practice are now available, particularly among apps that offer a full language-learning experience (e.g. *Duolingo*, *Busuu*).

A number of studies into the use of apps for language learning have been carried out. Yildiz (2012) found that using apps with young learners of English as a second language led to positive effects on vocabulary acquisition, phonological awareness and listening comprehension skills. A study with 33 undergraduate students of Spanish by Castañeda and Cho (2013) showed significant improvements in verb conjugation knowledge after using an app. Their participants also reported enjoyment of the gaming features of the app. Lys (2013) carried out a study of 13 university students of German. She found that the devices were suitable for speaking and listening activities at advanced level, and her students both felt comfortable using the devices and had the necessary competency to use them. Kim (2013) found improvements in listening comprehension among a group of Korean students and also reported positive attitudes towards the use of apps for this purpose, as did Khaddage and Lattemann (2013). Steel (2012) carried out a study of 134 language learners. Students reported that the features they liked best about using apps to support their learning outside class were flexibility, convenience, portability, and the ability to personalise their learning as well as using it on-the-go. The language areas that benefitted students most were vocabulary, reading and writing, grammar and translation activities. Steel found that many students used more than one app and valued the opportunities to engage with language learning outside the classroom. In a study with 85 distance learners of Spanish, Rosell-Aguilar (in press) also found that learners use apps mostly for vocabulary development, translation, and grammar practice. Students used apps often (44% used them at least once a day), mostly informally rather than in planned study sessions, and for relatively short periods of time. They liked the ability to practise specific areas, rapid access to information, ease of use, and gamification elements, but had concerns about usability and interface design, unreliability of content, lack

of grammar explanations, software errors, advertising, and poor feedback among others. All users reported that using apps improved their language skills to different degrees. Further studies have focused on specific skills for certain languages, such as learning non-western scripts (Rosell-Aguilar & Kan, 2015) with very positive results.

Although the use of apps can maximize the opportunities to engage in learning, the experience of learning on mobile devices can be highly fragmented and fraught with distractions (Kenning, 2007). One aspect of this fragmentation is the fact that users access their mobile devices for short amounts of time. This may affect learner choice of which app to use, as, for example, an app that requires listening or speaking may discourage use in a public place. Furthermore, Education apps have to fight for users' attention from strong competition from other apps within the device, such as games, and from pop-up notifications from social media, messaging or email, for example.

Most research into the evaluation of education apps has part focused on using one particular app within a concrete educational setting. This is no more useful than looking at a book as a single decontextualized learning solution. Apps are in many cases part of a suite of tools that a learner will use as part of their learning. This use of several apps to complement each other for a purpose is normally referred to as appsmashing.

The classification of the apps that can be used for language learning purposes can be approached from different angles. Previous classifications by Sweeney and Moore (2012), Rosenthal Tolisano (2012), and Schrock (2012) have mainly focused on learning skills, but these classifications did not clearly differentiate between those apps that have been developed for language learning purposes and those that have been developed for other purposes and can be of use to the language learner. A new taxonomy is proposed in section 3.

2.2 Evaluating language learning apps

A number of frameworks for the evaluation of education apps have been proposed (Walker, 2011; Schrock, 2011, 2013; Vincent, 2012; Peachey, 2013). Among the factors for the evaluation of effectiveness, a number of criteria are common to most frameworks. These include technical aspects, design, and whether the app is fit for purpose. The most frequently-mentioned criteria are curriculum connections / relevance and authenticity -whether targeted skills are practiced in an authentic format/problem-based learning environment. Other criteria include good navigation, support, accessibility, security, image and sound quality, usability, price, feedback, interaction, appropriateness of content, and instructions.

Typically, three approaches are used to evaluate software for CALL: checklists, methodological frameworks, and Second Language Acquisition (SLA)-based approaches (Levy and Stockwell, 2006). Jamieson, Chapelle and Preiss (2005) presented six criteria for evaluating CALL software which are in many ways still applicable today. They are: language learning potential, learner fit, meaning focus, authenticity, positive impact, and practicality. To these, others have added more detailed criteria (e.g. Hubbard, 2006). Many of these questions and criteria, however, looked at software (e.g. CD-ROMs) in the way it was provided at the time: as a single solution to be used extensively that had to be carefully selected considering price, platform, and necessary peripherals among other factors. In contrast with previous computer-based software, there is an enormous app market, cost is a fraction of what it used to be (which means apps can be downloaded, tested and deleted without huge investment loss), and the apps will be used on mobile devices rather than language labs or at a fixed location at a predetermined time. Most importantly, although some teachers may recommend the use of certain mobile apps or introduce them into their curriculum, it is mostly the users (autonomous learners in particular) who will make these choices independently.

Two frameworks have been proposed for evaluating language learning apps specifically: Sweeney and Moore (2012) listed the following criteria for evaluation: allowing personalization, visible progress indicators, covering relevant language, covering more than one skill, maximizing exposure to target language, appropriateness for the device (content, activity, interface), and encouraging learning behaviours which correspond with what we know about general mobile-enabled behaviour patterns (including social and gamification aspects). Rodríguez-Arancón, Arús and Calle (2013) presented a framework for evaluation of language learning apps covering the following criteria: cognitive value and pedagogic competence, content quality, capacity to generate learning, interactivity and adaptability, motivation, format and layout, usability, accessibility, visibility, and compatibility. This later framework is very detailed, presented with long descriptors in a rubric, which can be very helpful to the evaluator but adds complexity to the process. The descriptors of some of the criteria (format and layout, usability and accessibility in particular) overlap in ways that make them difficult to differentiate. Their criteria also miss out very relevant categories such as feedback, included in other frameworks.

Some authors (e.g. Walker, 2011) provide a minimum score they consider necessary for an app to be effective. Others suggest that the more criteria an app meets, the better it is (Vincent, 2012). Such statements are highly contentious. Since apps will serve different purposes for different learners depending on a number of circumstances such as the learner's language level or their personal learning preferences, to establish all the criteria as determining factors for the generic evaluation of an app could be misleading. Whilst some criteria are undoubtedly more crucial than others (e.g. some of the technical criteria – if the app does not work there is no possible learning value), one should not dismiss the potential of an app because it does not meet a certain criterion.

Another issue related worth mentioning in relation to evaluating apps is that most frameworks so far have been written by and for teachers and educators. It could be argued, however, that most app use will be outside formal learning opportunities and it is mainly autonomous users who need to evaluate the suitability of apps for their learning needs.

3 A taxonomy of mobile apps for language learning

The importance in education of establishing taxonomies is long established, dating back at least as far as what is known as Bloom's Taxonomy (Bloom, Engelheart, Furst, Hill & Krathwohl, 1956). Taxonomies are important and useful. As Krathwohl (2002) stated, Bloom believed that his taxonomy could serve, among other things, to provide a common language of reference, defining educational goals, and provide a panorama of educational possibilities (Krathwohl, 2002).

With the rise of new educational tools, such as apps, it is crucial that attempts are made to provide a similar taxonomy for the same reasons. Classifying apps into different types should help learners, teachers, and researchers to conceptualise and visualise the different varieties of apps available, which in turn can help to evaluate their potential, as this may depend on the type of app, intended audience and use (for formal tuition or autonomous learning, for example).

In Figure 1 a new classification of apps that can be used for language learning is presented, categorized in three groups according to whether they are primarily designed as language learning tools or not, and with a separate category for dictionaries and translators.

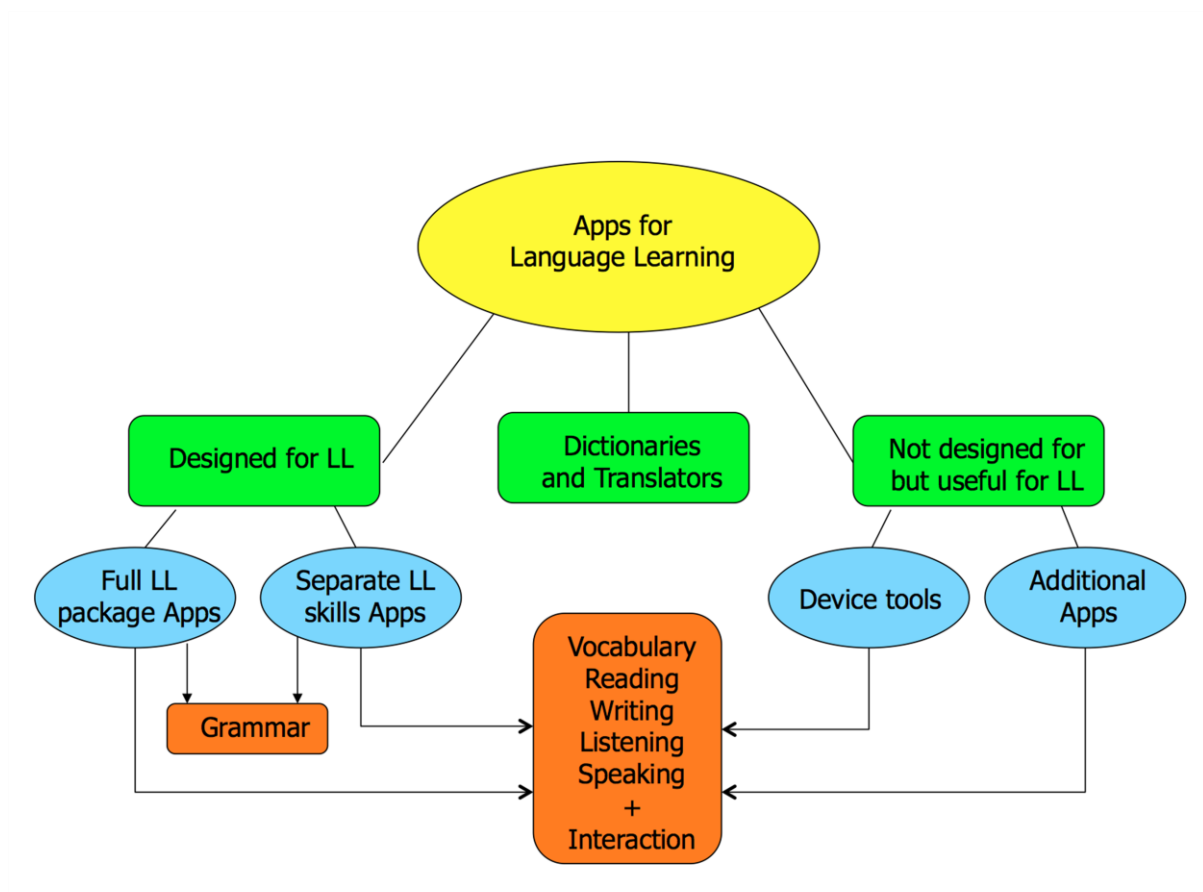


Figure 1: taxonomy of apps for language learning

3.1 Apps designed for language learning: the first group of these are apps that provide whole language learning packages: these apps are designed as full language learning solutions and offer a variety of exercises, grammatical explanations, and interaction with other students and native speakers as well as support from communities of learners. Some are mobile versions of previously-existing offerings. Most are free to download but many require a subscription to access full content. The most popular are *DuoLingo* and *Busuu*. Others include *Rosetta Stone*, *Speakeasy* and *Babbel*. Other apps aim to promote and keep alive lesser-known or endangered languages, such as the *Mixteco* app.

The second main groups of apps designed for language learning are those that offer activities to develop different areas of language such as grammar, vocabulary, reading, writing, listening and speaking, as presented in Table 1:

Area of language development	Description	Examples
Grammar	Grammar drills, some general and some more specific Verb conjugations	<i>French / Spanish grammar and practice series</i> <i>German gender trainer</i> <i>Bescherelle, Conjuverb, 501 Spanish Verbs.</i>
Vocabulary	Vocabulary drilling with images and sounds	<i>Learn German / French /Italian / Spanish series</i>
Reading	Literacy (mostly aimed at children) Graded readers	<i>Read me stories: learn to read</i> <i>Lire: French News reading and vocabulary</i>
Writing	Spelling practice apps Character writing apps Phonics	<i>Learn French Writing</i> <i>Spanish Spelling Tips</i> <i>Japanese-hiragana, Chinese First Steps</i> <i>Initial Code</i>
Listening	Texts in several languages with a read-along audio track	<i>BookBox</i>
Speaking	Pronunciation Phonetics	<i>iPronunciation</i> <i>MacMillan Sounds</i>
Interaction	Match language learners with partners or tutors for text, voice and / or video interaction either in real-time or asynchronously	<i>HelloTalk</i> <i>Tandem</i>

Table 1: taxonomy of apps designed for language learning

3.2 Apps not designed for language learning but useful to language learners. These may be device-native apps provided by default or additional apps that can be installed. The device-native tools that can aid the language learning process include language settings (although not an app per se, these can be changed so that menus and options, as well as apps

installed, will be in the target language); web browsers, which offer access to language learning web resources; multilingual text input (dictionary, grammar and auto-correct features can be set to the target language); speech-to-text tools, which can act as tools for testing pronunciation and to check spelling; communication tools such as email / messaging / telephone / video conferencing, which can provide opportunities for synchronous or asynchronous communication among learners, teacher-student, or with native speakers; the photo / video camera, which provide possibilities for creating content which can be the basis of or illustrate communicative exchanges; and even satellite navigators (if the language setting has been changed, directions will be provided in the target language).

Additional apps not native to the device that have uses for language teaching and learning are presented in Table 2:

Area of language development	Description	Examples
Vocabulary	Flashcard packages: although developed for any subject, learners can create their own sets with vocabulary, translations or conjugations to test their recall.	<i>Memrise, Quizlet</i>
Reading	Reading materials in the target language which cater for a variety of interests: e-books, comic books, news and magazine subscription apps.	<i>Kindle, Comic! Marvel Comics, BBC News, National Geographic</i>
Writing	Word processors with spell checkers Text sharing: Presentation apps Multimedia poster Storytelling Journal writing Blogging and microblogging	<i>Pages, Microsoft Word</i> <i>PowerPoint, Slideshare</i> <i>Phoster</i> <i>Our Story</i> <i>Day One</i> <i>Blogger, Wordpress, Twitter</i>
Listening	Podcast aggregators	<i>Podcast, iTunes U</i>

	Music streaming services and stores	<i>Spotify, iTunes, Soundcloud</i>
	TV programs and movie streaming and download services	<i>Netflix, iTunes, Amazon</i>
	Apps from national radio television broadcasters	<i>RTVE, France 24, RAI</i>
	Other video content	<i>YouTube, Vimeo, TED</i>
Speaking	Voice recorders	<i>QuickVoice</i>
	Video creation	<i>Vine, iMovie, YouTube</i>
Interaction	Communication tools in written, audio or video media	<i>Whatsapp, Line, Skype, FaceTime</i>
	Social media	<i>Facebook, Twitter</i>
	Social sharing networks for photographs, bookmarking	<i>Instagram, Flickr, Diigo, Pinterest</i>

Table 2: taxonomy of apps not designed for language learning but useful for language learners

In addition, other useful apps for the language learner and teacher include information resources (such as news apps), maps and geography (*Geomaster*) and geolocated information (*Aurasma, Wikitude*). Games can also be played in the target language (traditional games such as *Scrabble* or *Hangman*, or more current ones like *Clash of Clans*).

3.3 Dictionaries and translation apps: dictionary apps can be integrated into other apps, such as e-book readers, so that words can be looked up directly within the app. Some dictionary and phrasebook apps also include pronunciation examples. Translation apps offer machine translations with the option of entering text or speaking, and will pronounce the translation. Some examples are *Google Translate* and *iTranslate*. In this taxonomy they are classified separately as they are designed for both language learners and people who may not speak the language or be interested in learning it at all.

Although opposition to the use of translation apps has been raised by some language teachers, realistically these apps remain the first place where many language learners turn to when composing texts in the target language. The machine translation algorithms have improved vastly in recent years, but translations can be erroneous, especially when words are looked up decontextualized, and learners should be encouraged to evaluate their output for possible errors or editorial needs, as they would with a dictionary.

4 A framework for evaluating language learning apps

Apps can provide a vast array of affordances for language learners and teachers, but aside from highlighting their potential, and given the large number of apps of varying quality available to download, it is essential that learners, teachers and researchers have the tools to evaluate them. The framework for app evaluation proposed here is based on some categories from frameworks presented in section 2 as well as on SLA principles of task design, presented in a simple format for ease of use by both learners and educators.

When designing activities for language learning, cognitive and interactionist SLA principles advocate Task-based Language Teaching based on concepts including noticing, negotiation of meaning, learning by doing, focus on form and collaborative learning (Skehan, 2003; Doughty & Long, 2003). From SLA literature we surmise that language learning tasks should be interactive and include reporting back of the communicative outcome (Skehan 2003), collaborative, interesting, rewarding, and challenging (Meskill, 1999), meaningful and engaging rather than repetitive or stressful (Oxford, 1990), provide opportunities to produce target language (Chapelle, 1998), and make use of authentic materials (Little, 1997). Furthermore, it is known that learners' performance improves if they feel in command of the situation, and if they are familiar with their environment (Oxford, 1990) so the usability of the design of an app - how easy to learn and use it is - is very important.

The new framework proposed here is divided into four primary categories: technology, user experience, pedagogy and subject specific (in this case language learning), each with a number of criteria. The evaluation framework is presented in Table 3 as a list of questions for use by learners and educators alike to help them decide whether an app meets their learning and teaching needs.

Language learning	Pedagogy
<ul style="list-style-type: none"> ▪ Reading: does the app provide texts in the target language? ▪ Listening: does the app provide audio in the target language? ▪ Writing: does the app offer opportunities to write in the target language? ▪ Speaking: does the app offer opportunities to speak in the target language? ▪ Vocabulary: does the app offer specific activities for vocabulary acquisition? ▪ Grammar: does the app offer specific activities for grammar practise? ▪ Pronunciation and intonation: does the app offer specific activities for pronunciation and intonation? ▪ Cultural information: does the app include information about customs and traditions in the areas where the language is spoken? ▪ Use of visual content: are images and videos stereotypical or stock images? Do they represent the diversity of the areas where the language is spoken? ▪ Language varieties: does the app include different regional or national varieties of the language? 	<ul style="list-style-type: none"> ▪ Description: does the app store description match what the app does? ▪ Teaching: does the app present, explain or model language or does it just test it? ▪ Progress: does the app allow the user to track progress or see previous attempts? ▪ Scaffolding: do activities in the app progress in difficulty in a way that supports the learner? ▪ Feedback: does the app provide feedback? Is it just right/wrong or meaningful explanations? ▪ Quality of content: does the content have any errors / omissions? ▪ Use of media: does the app make use of sound, images and video in a meaningful way? ▪ Differentiation: does the app offer different levels depending on ability? Can these be accessed directly? ▪ Engagement: does the app keep the user interested or are activities repetitive?
User experience	Technology
<ul style="list-style-type: none"> ▪ Interaction: does the app allow users to interact with each other? ▪ Interactivity: is engagement with the app content active or passive? ▪ Sharing: does the app allow or encourage sharing content? 	<ul style="list-style-type: none"> ▪ Interface: is the interface clear and uncluttered? ▪ Navigation: is the app intuitive to navigate, with clear menus and options? ▪ Instructions: does the app offer instructions on how to use it?

<ul style="list-style-type: none"> ▪ Badging: does the app provide recognition that can be shared on social media? ▪ Price: does the user need to pay to download the app? Is there a “lite” version? Does it offer in-app purchases? ▪ Registration: does the app require the user to register? ▪ Advertising: does the app include pop-up ads? Are these distracting? 	<ul style="list-style-type: none"> ▪ Stability: does the app freeze or crash? ▪ Gamification: does the app have game-like features to increase engagement? ▪ Support: does the app have a help section? ▪ Offline work: does the app require an internet connection to work?
---	--

Table 3: Framework for Language Learning app evaluation

There is a degree of overlap between the criteria, and some of them apply to more than one of the four main categories. For example, Feedback could apply to technology (in terms of how it is presented), pedagogy (how it relates to teaching), language learning (the quality of the feedback) and user experience (how well the feedback fits in the learning process, where it appears, how it can be accessed). In Figure 2 the framework is presented in visual form, although this is somewhat subjective and limited by visual representation.

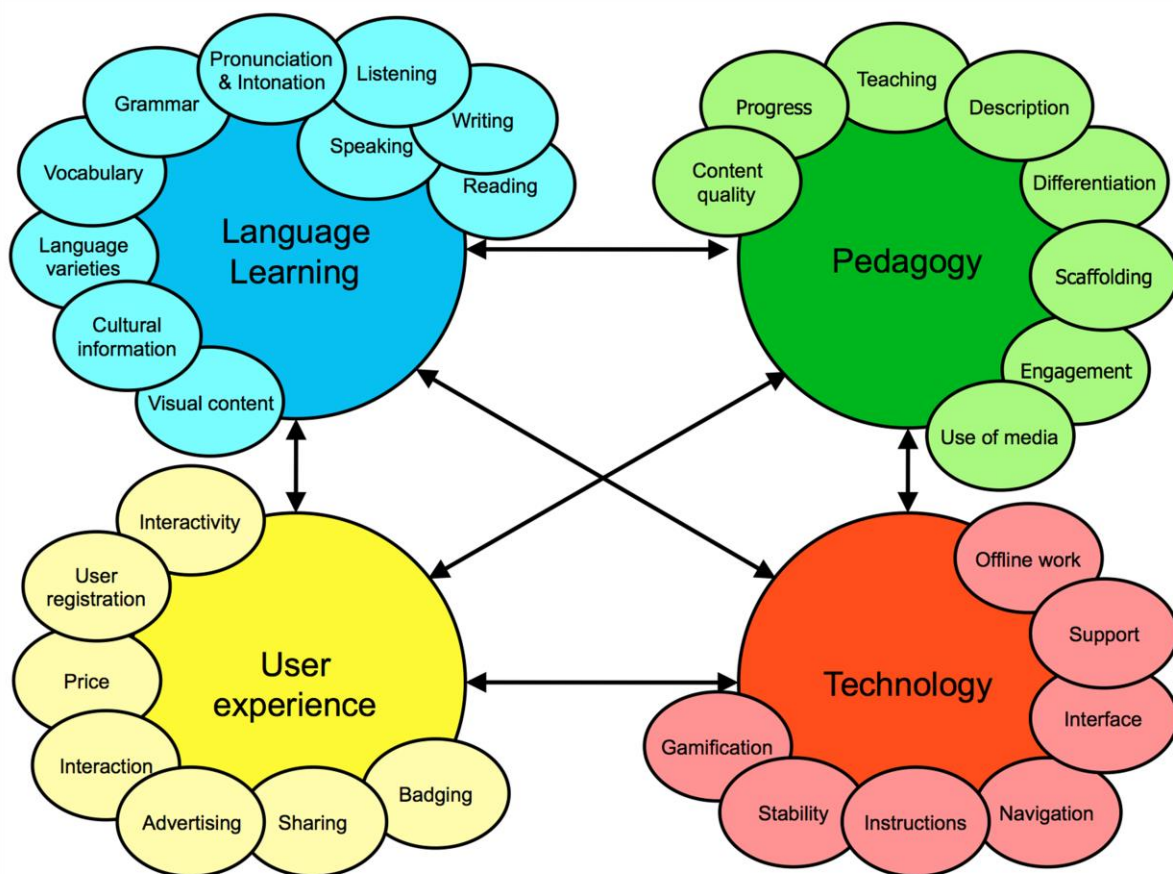


Figure 2: categories and criteria for the evaluation of mobile learning apps

This list of questions does not offer a rubric with detailed descriptions of each criterion for two reasons: to keep the questions clear and uncluttered, and because the aim of the questions is not to award a mark or value to each question, but for the questions to act as a reflection tool for both learners and teachers, as well as app developers and researchers. There is no indication in the framework about how many of the criteria an app needs to meet to be considered apt for language teaching or learning because different learners may find an app useful or not depending among other reasons on their purpose, learning preferences, location and personal circumstances. In addition, some criteria will only apply to an app depending on what it is supposed to do. There would be no gain, for example, in appraising a vocabulary app negatively for not offering speaking practice, although a more comprehensive evaluation, with positive appraisals for a higher number of the criteria, would be expected for apps that claim to offer a full language learning experience.

It is important to stress that this evaluation framework applies to commercially available self-contained apps that can be installed on devices such as smartphones and tablets, and not to all resources that can be accessed through such devices, such as ebooks or web resources.

An early version of this framework was tested in a workshop in Ireland with a group of 18 language teachers in October 2014. After a presentation of the framework, participants were asked to evaluate the apps they use for language teaching using the criteria in the framework. Participants provided oral feedback in a short focus group activity at the end of the workshop. All participants were positive about the use of the criteria and supported that the criteria helped them shape their own evaluation of language learning apps. It was mentioned that, since most students own smartphones and / or tablets, it would be a worthwhile activity to spend time in class presenting the framework to language learners to enable them to make better-informed decisions about which apps are suitable for them depending on the curriculum as well as their own learning preferences and needs. Suggestions for changes to the framework included revising the descriptions for clarity, and separating some of the categories. The original framework only had two main categories (Pedagogy and Technology) and, upon further reflection after the workshop, the four-category model was created.

In addition, a second workshop with a different group of 26 language teachers took place in Cyprus in November 2015. Following a similar format, the feedback this time focused on the Language Learning category, which some the teachers felt was too abstract. Based on this feedback, that section was rewritten to change criteria that referred to SLA and MALL theories for the current criteria presented, thus making it clearer to use and dispensing with the need for users to be aware of current SLA trends when utilising the framework.

5 Further research

Although the experience of mobile device use in the classroom has been well documented, the amount of research examining how learners engage in mobile learning outside the classroom is much smaller (Stockwell, 2013). There is much potential for research in the field of mobile apps for language learning, including the following:

- App design and quality: Can apps offer true language learning solutions? What do language learning apps offer to the learner that other more traditional methods do not (and vice versa)?
- Users: As part of research into the use of apps, questions that should be asked include: Who uses language learning mobile apps? Why? Where? How? What do they think about learning with apps?
- Appsmashing: how apps are used in combination with other resources remains an interesting topic still under-researched.
- Normalization: at what point do we consider the use of smartphones and tablets normalized (Bax, 2003) to the point that they are fully integrated into learning activity? Can we assume learners own such devices and have the competencies to know how to use them, select appropriate resources, and utilize them when and where they are best served by them?
- Attainment: although the potential for learning is there, further research is needed on learning outcomes.

Some of this research, in particular research into actual gains in language proficiency, will be difficult to carry out as learners tend to use apps in combination with other apps or to

supplement other forms of learning, formal or informal, which makes causality difficult to prove.

6 Conclusion

This paper has provided an evaluation framework for language learning mobile apps, but has not evaluated the apps themselves. A proposal to make this framework available on a dedicated website for language apps evaluation is currently being considered.

Developments in mobile app software are fast and it is difficult to foresee what direction software and hardware will take next. Wearable technologies will undoubtedly provide new affordances for learning, but whether they succeed in penetrating the mainstream (or not, as the Google Glass initiative has proven so far) and their effect on mobile learning will be an interesting development to ‘watch’.

Acknowledgements:

I would like to thank Anna Comas-Quinn for her useful comments on an earlier draft of this paper.

In memoriam: I would like to dedicate this paper to Lesley Shield, a true CALL pioneer and wonderful colleague, from whom I learnt a lot about the evaluation of CALL resources among other things.

References:

- Ally, M. (2005). Using learning theories to design instruction for mobile learning devices. *Mobile learning anytime everywhere*, 5--8.
- Bax, S. (2003). CALL—past, present and future. *System*, 31(1), 13--28.
- Bloom, B. S., Engelhart, M. D., Hill, W. H., & Furst, E. J. Krathwohl (1956). *Taxonomy of educational objectives. Handbook I: Cognitive domain*. New York: McKay.
- Burston, J. (2014). The Reality of MALL: Still on the Fringes. *CALICO Journal*, 31(1), 103--125.
- Castañeda, D. & Cho, M.-H. (2013). *Using Smartphone Applications to Improve Students' Verb Conjugations in Spanish: An Exploratory Study*. Paper presented at the CALICO Conference, Navigating the Complexities of Language Learning in the Digital Age, Hawaii, US, 24th May 2013.
- Chapelle, C. (1998). Multimedia CALL: lessons to be learned from research on instructed SLA. *Language Learning and Technology*, Vol. 2, No. 1, 22--34. Retrieved from <http://llt.msu.edu/vol2num1/article1/index.html>
- Colpaert, J. (2006). Pedagogy-driven design for online language teaching and learning. *CALICO journal*, 477--497.
- Domínguez, A., Saenz-de-Navarrete, J., De-Marcos, L., Fernández-Sanz, L., Pagés, C., & Martínez-Herráiz, J. J. (2013). Gamifying learning experiences: Practical implications and outcomes. *Computers & Education*, 63, 380--392.
- Doughty, C. & Long, M. H. (2003). Optimal Psycholinguistic Environments for Distance FL Learning. *Language Learning and Technology*, 7 (3), 50--80. Retrieved from <http://llt.msu.edu/vol7num3/doughty/default.html>
- Godwin-Jones, R. (2011). Emerging technologies: Mobile apps for language learning. *Language Learning & Technology*, 15(2), 2--11.
- Hubbard, P. (2006). Evaluating CALL Software. Chapter 13 from L. Ducate and N. Arnold (eds) 2006. *Calling on CALL: From Theory and Research to New Directions in Foreign Language Teaching*. San Marcos: CALICO.
- Jamieson, J., Chapelle, C., & Preiss, S. (2005). CALL Evaluation by developers, a teacher, and students. *CALICO Journal*, 23 (1), 93--138.
- Kenning, M. M. (2007). *ICT and language learning: From print to the mobile phone*. Palgrave Macmillan.
- Khaddage, F., Lattemann C., (2013). The Future of Mobile Apps for Teaching and Learning. In Berge, Z. L., & Muilenburg, L. (Eds.). *Handbook of Mobile Education*. Routledge, 119--128.
- Kim, H., & Kwon, Y. (2012). Exploring smartphone applications for effective Mobile-Assisted Language Learning. *Multimedia-Assisted Language Learning*, 16(1), 31--57.
- Kim, H. S. (2013). Emerging Mobile Apps to Improve English Listening Skills. *Multimedia-Assisted Language Learning*, 16(2), 11--30.

- Krathwohl, D. R. (2002). A revision of Bloom's taxonomy: An overview. *Theory into practice*, 41(4), 212--218. Retrieved from http://www.unco.edu/cetl/sir/stating_outcome/documents/Krathwohl.pdf
- Levy, M., & Stockwell, G. (2006). *CALL dimensions: Options and issues in computer-assisted language learning*. Mahwah, NJ: Lawrence Erlbaum.
- Little, D. (1997). Responding authentically to authentic texts: a problem for self-access language learning? In Benson, P. & Volley, P. (eds.), *Autonomy and Independence in Language Learning*: Longman, 225--236.
- Lys, F. (2013). The development of advanced learner oral proficiency using iPads. *Language Learning & Technology*, 17(3), 94--116. Retrieved from <http://llt.msu.edu/issues/october2013/lys.pdf>
- Meskill, C. (1999). Computers as Tools for Sociocollaborative Language Learning. In K. Cameron (Ed) *CALL: Media, Design and Applications*. The Netherlands: Swets & Zeitlinger. Retrieved from <http://www.albany.edu/etap/Faculty/CarlaMeskill/publication/cameron.pdf>
- Oxford, R. (1990). *Language learning strategies: What every teacher should know*. New York. Newbury House.
- Peachey, N. (2013). *Criteria for Evaluating Web Tools and Apps*. Retrieved from: <http://nikpeachey.blogspot.co.uk/2013/04/criteria-for-evaluating-web-tools-and.html>
- Rodríguez-Arancón, P., Arús, J., & Calle, C. (2013). The use of current mobile learning applications in EFL. *Procedia-Social and Behavioral Sciences*, 103, 1189--1196.
- Rosell-Aguilar, F. (2009). Podcasting for language learning: re-examining the potential. In L. Lomicka & G. Lord (Eds.), *The Next Generation: social networking and online collaboration in foreign language learning*, (pp. 13--34), San Marcos, TX: Calico.
- Rosell-Aguilar, F. (in press). User evaluation of language learning mobile applications: a case study with learners of Spanish. In A. Palalas & M. Ally (Eds), *The International Handbook of Mobile-Assisted Language Learning*. China Central Radio & TV University Press.
- Rosell-Aguilar, F. & Kan, Q. (2015). Design and user evaluation of a mobile app to teach Chinese characters. *JaltCALL Journal*, 11(1), 19--40.
- Rosenthal Tolisano, S. (2012). *iPad App Evaluation for the Classroom*. Retrieved from <http://langwitches.org/blog/wp-content/uploads/2012/06/iPadAppEvaluation.pdf>
- Shrock (2012). *Bloomin' Apps*. Retrieved from <http://www.schrockguide.net/bloomin-apps.html>
- Skehan, P. (2003). Focus on Form, Tasks, and Technology. *Computer Assisted Language Learning*, 16 (5), 391--411.
- Steel, C. (2012). Fitting learning into life: Language students' perspectives on benefits of using mobile apps. In: M. Brown, M. Hartnett and T. Stewart, *ascilite2012 Conference proceedings*. Wellington, New Zealand, (875--880).

Sweeney, P., & Moore, C. (2012). Mobile Apps for Learning Vocabulary: Categories, Evaluation and Design Criteria for Teachers and Developers. *International Journal of Computer-Assisted Language Learning and Teaching (IJCALLT)*, 2(4), 1--16.

Vincent, T. (2012). *Ways to Evaluate Educational Apps*. Retrieved from <http://learninginhand.com/blog/ways-to-evaluate-educational-apps.html>

Walker, H. (2011). Evaluating the effectiveness of apps for mobile devices. *Journal of Special Education Technology*, 26(4), 59--63.

Yildiz, S. (2012, June). *Use of iPad applications to introduce English as a foreign language to young Turkish learners*. Paper presented at the 2012 CALICO Conference, University of Notre Dame, South Bend, IN.